SECTION 6

QUESTION 5: What supporting data are required for the calibration/validation of the model on the spatial and temporal scales necessary to address the principal need for the model (as defined above)? What supporting data are required to achieve the necessary level of process resolution in the model?

6.1 SUPPORTING DATA FOR HSPF

Supporting data needed for calibration and validation of the watershed model (HSPF) are described in the MFD (page 5-2) and QAPP (page 4-17). These data include, but are not limited to: meteorological inputs, watershed characteristics, in-stream flows and sediment loads, and nutrient parameters. Adequate data sets exist to calibrate the hydrologic and solids transport submodels of the watershed modeling effort. However, calibration of the PCB transport submodel may be problematic because of a lack of data characterizing two potential PCB sources: 1) groundwater and 2) surface runoff. As discussed in Section 2.3, it would be better if watershed PCB loadings were not developed using HSPF, but rather established from site-specific data. In addition, it is unclear as to whether sufficient data exist to calibrate and validate the nutrient transport submodel.

6.2 SUPPORTING DATA FOR EFDC

Supporting data needed for calibration and validation of the hydrodynamic, sediment transport, and abiotic PCB models, using EFDC, are described in the MFD (page 5-3) and QAPP (page 4-28). These data include, but are not limited to: water surface elevations (stage height), current velocities, TSS concentrations, water column PCB concentrations under low-flow and high-flow conditions, distributions of sediment types, bathymetric data, historical and current sediment PCB concentrations and distribution, and high-resolution cores to estimate deposition rates in impoundments. For the most part, sufficient data are available to calibrate all three submodels within EFDC. These data include, but are not limited to, water column monitoring data from the 1990s (including routine and storm sampling), sediment PCB concentrations measured between 1980 and 2000, and high-resolution sediment cores from impoundments to estimate deposition rates.

However, as discussed in Section 2.2.8, to the extent that sediment bed load transport will be modeled, some direct measurements of sediment bed load will be necessary to calibrate that component of the model. Such measurements should be made at a number of stations and river flow rates, but are particularly important in Reach 5A, including the upstream model boundary.

6.3 SUPPORTING DATA FOR AQUATOX

The MFD states that for the ecosystem variables in AQUATOX, little calibration will be necessary (MFD page 5-7). However, the requirement for calibration to site data cannot be determined *a priori*, but will depend on the match between the model and the data. Thus it is inappropriate to state at the outset that little calibration will be needed for certain parts of the model.

For some AQUATOX variables, particularly species biomasses, site data will be limited. As mentioned above (Sections 2.5 and 3.4), AQUATOX varies the diet of species simulated based on their calculated biomasses. This affects the pathway of PCB transfer from sediments and water to the biota. Given the variability in species density throughout the year and among years and the lack of site-specific data on such species densities in the Housatonic River, the biomass calculations cannot be adequately constrained. Therefore, as discussed previously, those biomass calculations should not be used in the model; rather, the diet inputs to the model should be estimated directly, based on available information, and uncertainties in the calculations should be assessed through model sensitivity and uncertainty analyses.

The MFD also indicates that AQUATOX will be validated by performing a model simulation for the period from 1980 to 2000 (MFD page 5-8). However, as noted in the MFD, the historical PCB data are limited, thus limiting the value of validating the PCB bioaccumulation model based on those data. PCB data for fish in the model reach are available for recent years (including the PCB fish data collected by USEPA and GE in 1998 and the young-of-the-year fish data collected by GE in 1994, 1996, 1998, and 2000), but for prior years PCB fish data from this reach are available only from 1982, with limited supplemental sampling in Woods Pond in 1990. Water column PCB data are available within the study area only since 1989, and the data from that year and the early 1990s have high detection limits and are limited in number. In these circumstances, while a historical simulation would be of interest, it will not provide a strong additional constraint on the PCB fate and bioaccumulation models.

The MFD notes further that the uncertainty associated with AQUATOX will be assessed using a probabilistic Monte Carlo approach (MFD pages 5-12 to 5-14). This approach requires characterization of the shape of the distribution of each key parameter, estimation of the parameters of each distribution, (*i.e.,* mean and standard deviation for a normally distributed variable) and estimation of correlations among parameters. However, the site-specific data are insufficient to estimate all of these parameters. Thus, the realism of the distributions of computed PCB concentrations will not be known, compromising the utility of the results. Therefore, Monte Carlo simulations should not be performed.

The key question associated with any model is: can alternative sets of parameters produce acceptable calibrations, while producing different projections that result in different management decisions? This question can be approached directly through the development of one or a few alternative calibrations. Some or all projections would then be performed using both the best calibration and the alternative calibration(s). In this way, the impact of model uncertainty on management decisions can be assessed. An example of this strategy can be found in QEA (1999).

6.4 USE OF MODEL CALIBRATION METRICS

The QAPP proposes use of certain metrics to govern model calibration (QAPP Table 4-4), following a strategy outlined in a Technical Memorandum from Wisconsin Department of

Natural Resources concerning the development of fate, transport and bioaccumulation models for PCBs in the Fox River and Green Bay. These metrics rely heavily on specific numerical criteria for model reliability, which are given as limits on the difference between model predictions and data. For example, sediment PCB concentrations must be within 50% of the data (QAPP, Table 4-4). This does not represent the best way in which model reliability can be assessed, because:

- Uncertainty in the data may make such specific criteria inappropriate. For example, computed sediment PCB concentrations may lie within the uncertainty range of the data, which may be specified as 95% confidence limits on the mean PCB concentration, while the model/data difference may be greater than 50%. In this case, the model would be consistent with the data, but would be deemed unsatisfactory according to the proposed metric.

- The model may exhibit a consistent bias with respect to the data, yet remain within the specified limits and be deemed acceptable under the proposed metric.

- The model may exhibit a temporal or spatial trend that is inconsistent with the data, yet remain within the specified limits and be deemed acceptable under the proposed metric.

- These metrics do not permit incorporation of the relative reliability of different data sets. Data sets differ in reliability, for example because of the numbers of values or differences in laboratory technique. Model/data comparisons must take such differences into account, because it may be impossible for a model to meet the criteria for all data simultaneously. However, specifying precise limits for model/data differences beforehand precludes consideration of these factors.

Therefore, precise limits on model/data differences should not be specified in the QAPP. Specific tests of the model should be specified, but should incorporate both uncertainty in the data and uncertainty in values used for key model parameters. An alternative set of model performance guidelines is as follows:

- Do the model results generally fall within the error bars of the data?
- Are model predictions generally unbiased with respect to the means of the observed data?
- Are temporal and spatial trends computed by the model consistent with trends observed in the data?

6.5 GEOMORPHOLOGICAL STUDY

USEPA proposes to conduct a geomorphological study of the River using the Rosgen (1994) method (MFD page 6-4), presumably to moderate the results of the PCB fate modeling. In this respect, this study is as important as any of the model calibration and validation procedures. The stated purpose of the Rosgen study is to determine whether man-made influences are contributing to channel instability within the study area. Data collection programs to measure bank erosion rates and bed elevation changes will be included in this investigation. The MFD does not provide any details as to how the study will be conducted or related to the modeling study. The MFD should provide additional details and information regarding the geomorphological study. Specific questions that need to be address include:

- How does this study relate to the overall goals of the modeling study?
- What exactly will USEPA do to determine if the channel is unstable?
- The study area needs to be compared to a river system within the same physiographic region that is unimpacted. What river will be used for comparison? What criteria will be used to determine whether or not that river is unimpacted?
- How will the Rosgen (1994) method be applied to the study area?
- Will any efforts be made by USEPA to determine whether or not the results of the modeling and geomorphological studies are consistent? If results from these studies are inconsistent, which approach will be used to evaluate remedial alternatives regarding bed stability issues?
- How will data collected during this study be used?